

Middlesex University Research Repository

An open access repository of

Middlesex University research

<http://eprints.mdx.ac.uk>

Li, Simone, Xu, Kai ORCID logoORCID: <https://orcid.org/0000-0003-2242-5440> and Wilkins, Marc (2010) Visualisation of the complexome. In: Human Proteome Organisation 9th Annual World Congress, 19th-23rd September 2010., Sydney Convention and Exhibition centre. .
[Conference or Workshop Item]

This version is available at: <https://eprints.mdx.ac.uk/7041/>

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author's name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: <http://eprints.mdx.ac.uk/policies.html#copy>

Simone S. Li¹, Kai Xu² and Marc R. Wilkins¹

¹Systems Biology Initiative and School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia

²School of Engineering and Information Sciences, Middlesex University, London, UK

Introduction

- Proteins and other molecules interact with each other to form macromolecular complexes, which are the functional units inside the cell. Together, these make up the complexome.
- To date, there are 4 large-scale datasets for protein complexes in *Saccharomyces cerevisiae*. These were established by affinity purification and mass spectrometry.
- Of these, one featured protein sharing between complexes (Gavin et al.¹). In this study, they proposed a 'core-module-attachment' model. Core proteins are always present within a complex. Attachment proteins are single proteins that are sometimes but not always in a complex. Modules are groups of two or more strongly associated attachment proteins present in more than one complex.
- The aim of this study is to visualise the complexome as a network. This should provide a unique 'high-level' and biologically relevant view of the interactome.

The Complexome as a Network of Protein Complexes

- The dataset by Gavin et al¹ was used as a basis for this study.
- In our model, protein complexes are connected if they share one or more proteins [Fig. 1]. The lines between complexes represent the proteins in common.

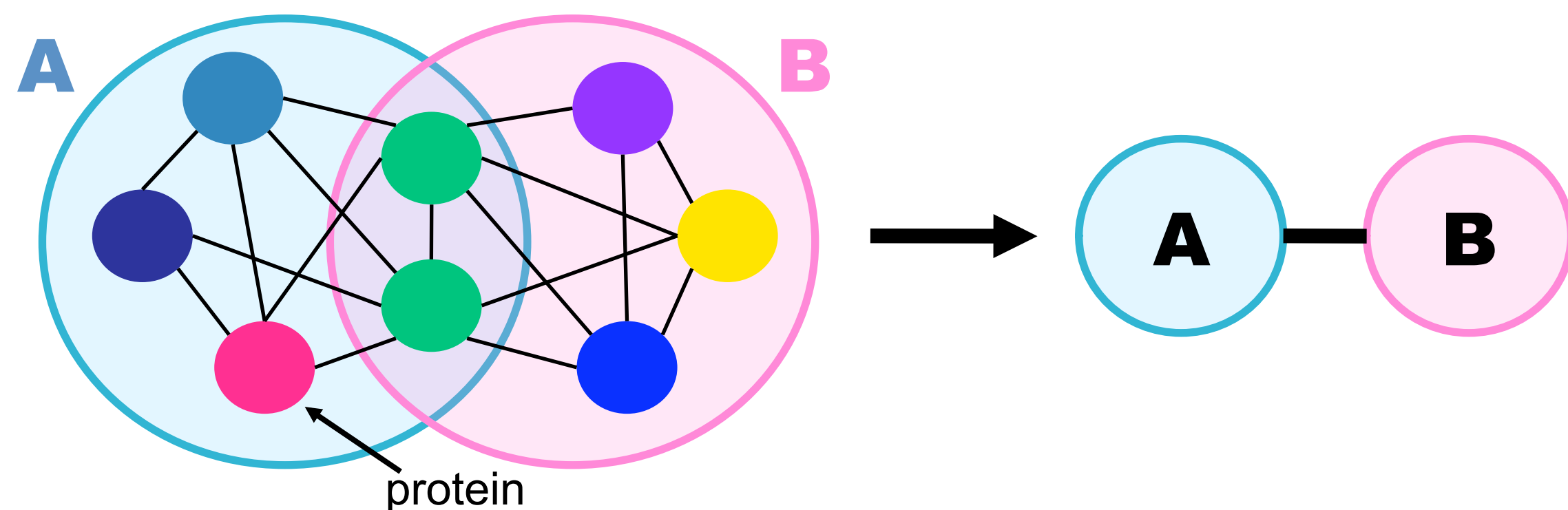


Figure 1: From interactome to complexome. In this example, two protein complexes A and B share 2 proteins. In the interactome, this is visualised as a highly interconnected cluster, and may be mistaken as a single complex. In the complexome, A and B are collapsed into single nodes, each representing a complex. The line drawn between them represents the 2 proteins they share.

- The GEOMI² platform was modified to visualise the complexome [Fig. 2]. This can be downloaded at <http://www.systemsbiology.org.au>.

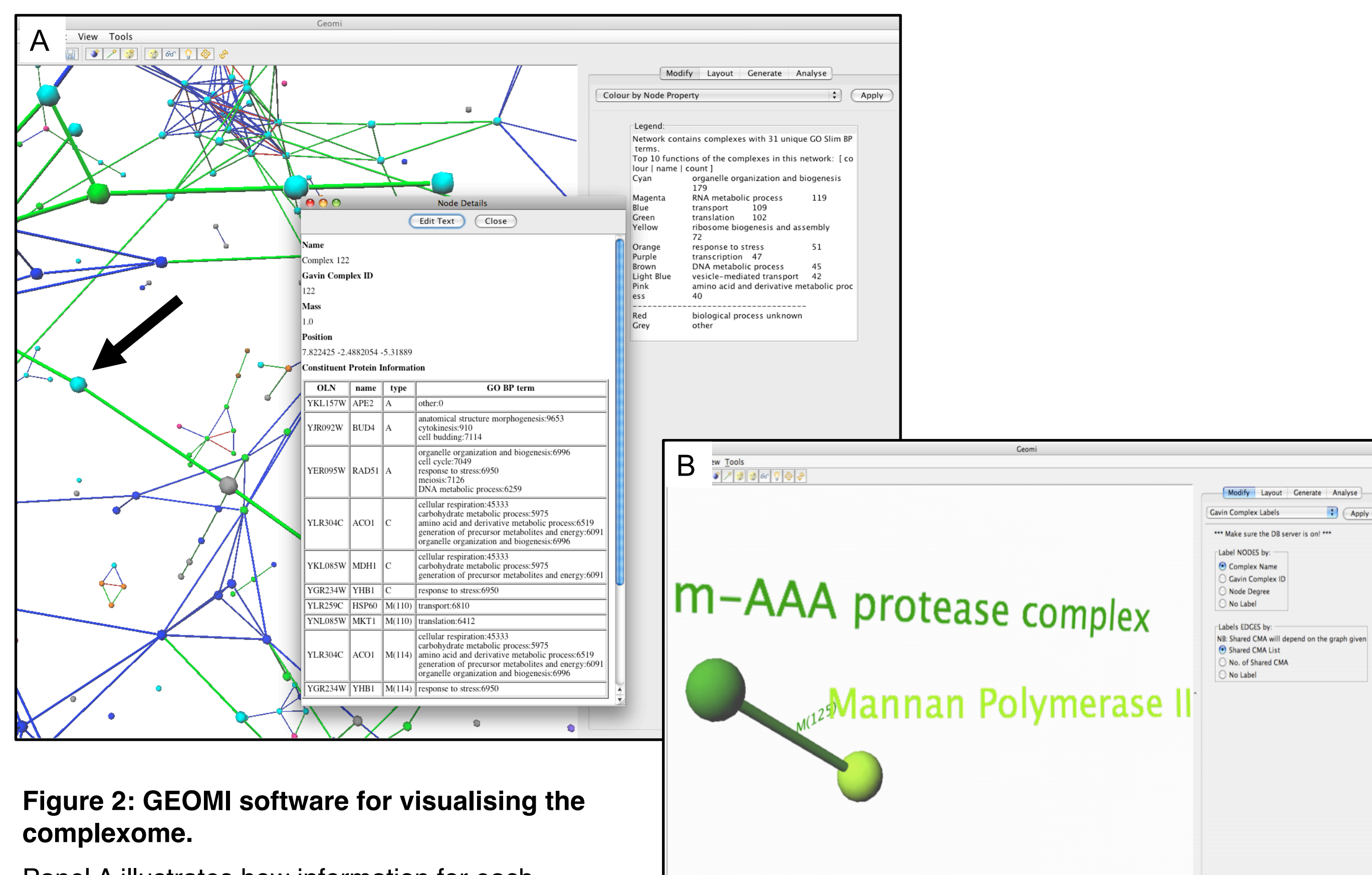


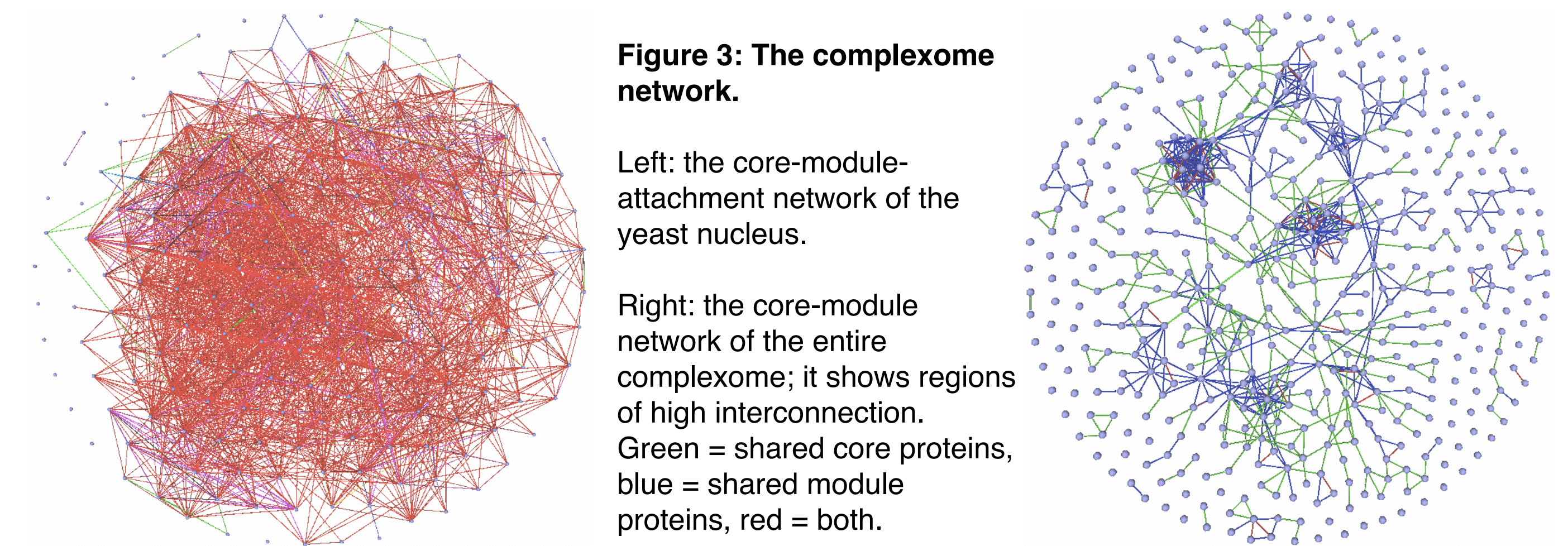
Figure 2: GEOMI software for visualising the complexome.

Panel A illustrates how information for each complex can be retrieved and displayed from an underlying relational database. In this case, complex 122 (arrow) has been selected via a right mouse button. The types of proteins in each complex (core, attachment, module) can be seen.

Panel B shows two complexes, joined together by module 125. The control panel (top right) has many options for labelling nodes and edges.

Structure and Topology of the Complexome Network

- The complexome network was constructed in two ways. Networks including links from shared core, module or attachment proteins were dense and uninterpretable [Fig. 3]. Networks whereby complexes were linked from shared core or module proteins showed some regions of high interconnectivity and other areas with no connections. This network is the CM (core-module) network.



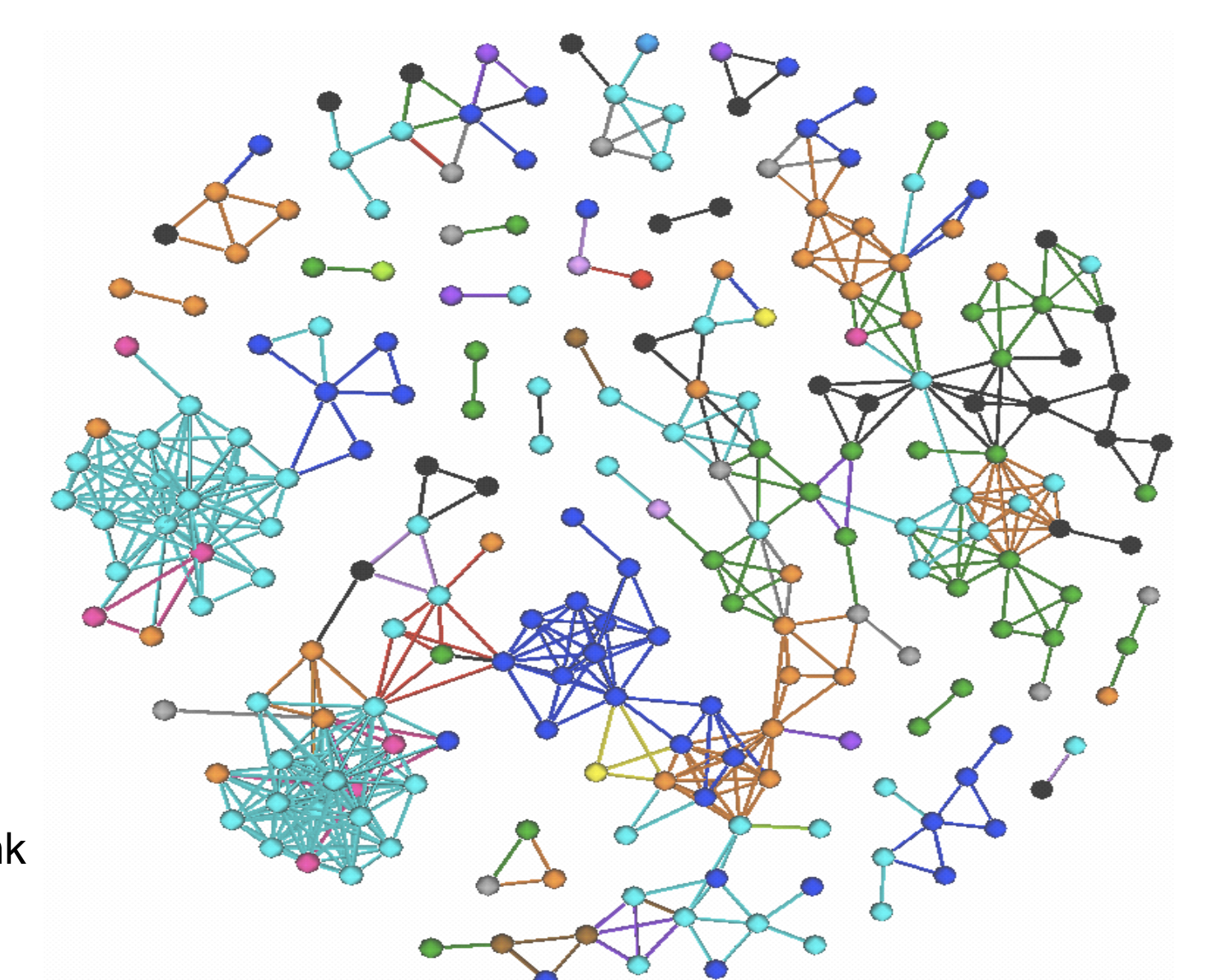
- The CM network shows properties of scale-free networks. It has a degree exponent γ of 1.9, where scale-free networks have exponents of 2 to 3. It also has a clustering coefficient which is independent of node degree; a further scale-free property.

Functional Relationships of Interacting Complexes

- Gene ontology was used to investigate the functional relationships of interacting complexes.
- The biological processes of constituent proteins from each complex were tallied to assign a high-level function to each complex. This was then co-visualised with the CM network [Fig. 4].
- Many interacting complexes share common functions, suggesting the CM network is biologically relevant.

Figure 4: The complexome network, coloured by biological process. Node colour and line colours can be different, if shared proteins are of different process to those that are not.

Cyan – organelle organisation, Blue – RNA metabolic process, Orange – translation, Green – transport, Yellow – transcription, Pink – ribosome biogenesis, Black – other, Red – unknown. Singletons not shown.



Interacting Complexes Reflect Higher Order of the Cell

- The higher order of the cell is reflected in interacting complexes [Fig. 5]. This suggests that 'guilt by association' will be useful to understand the function of novel complexes.

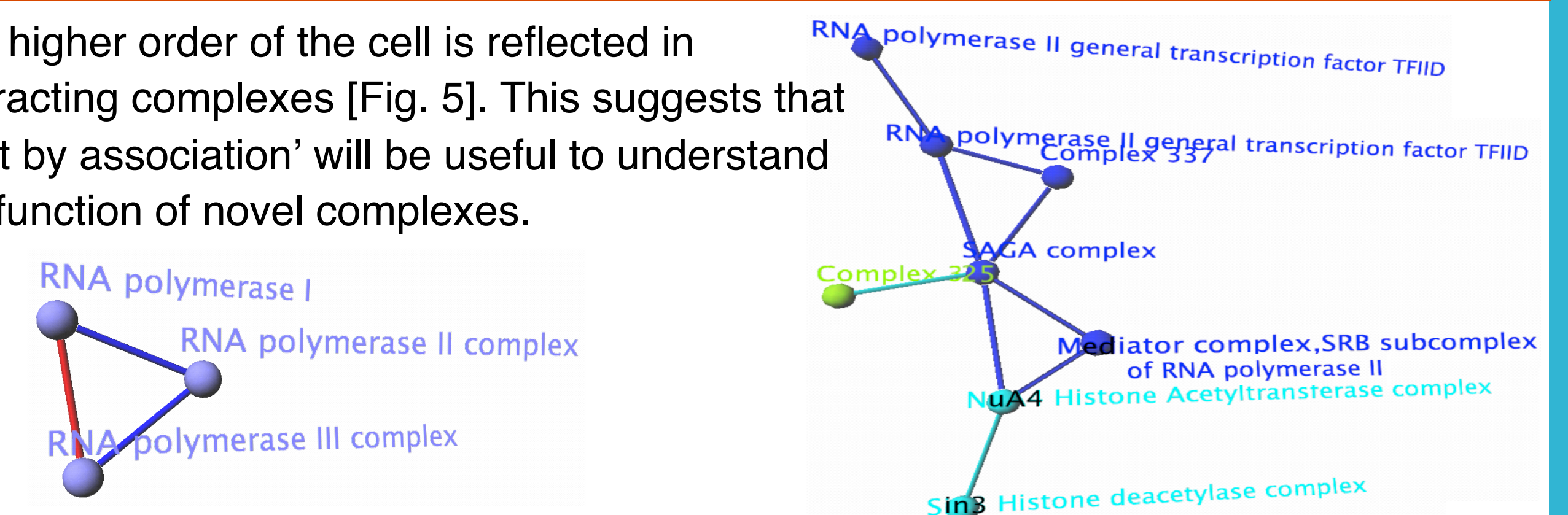


Figure 5: Detail of the complexome network. The RNA polymerase complexes I to III show interconnection, as do the SAGA, mediator and histone acetylases and deacetylases. Complexes 325 and 337 are likely to be involved in related processes.

Acknowledgements and References

We acknowledge funding from the NSW Office for Science and Medical Research (OSMR), EIF Super Science, the Australian Research Council and UNSW. We also acknowledge the assistance of NICTA.

¹Gavin AC, et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature*, **440**, 631-636.

²Ho E., Webber R. and Wilkins M. (2008). *Journal of Proteome Research*, **7**, 104-12.